DOCUMENT RESUME

ED 295 970 TM 011 761

AUTHOR Davis, Todd M.; And Others
TITLE An Investigation of Speededness in a Placement Test: A Call for Local Standards.
PUB DATE Apr 88
NOTE 29p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *College Applicants; College Entrance Examinations; Disadvantaged; Higher Education; Remedial Instruction; Standards; *Student Placement; Test Format; Test Length; *Timed Tests
IDENTIFIERS *Academic Assessment and Placement Program; *Speededness (Tests); Tennessee

ABSTRACT

The effect of time limits on the completion rate of 8,290 students taking the reading comprehension section of the Academic Assessment and Placement Program (AAPP) was studied. The AAPP is a Tennessee configuration of items drawn from the College Board's multiple assessment programs and services item pool. It is a battery of tests used for diagnostic placement of Tennessee students into remedial or developmental courses. Students under 21 years old with an American College Tests (ACT) Composite Score of 15 or below and students 21 years old or older with no ACT test on file are required to take the AAPP tests in English, mathematics, reading comprehension, and logical relationships. For the study population as a whole, and by sex and age, no substantial effects of time limits were observed. Differential completion rates by race were noted at all 20 Tennessee State Board of Regents institutions evaluated, with whites having higher completion rates than blacks. Such differences should become a priority for study. Seven graphs are presented. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED295970

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
JERRY N. BOONE
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

AN INVESTIGATION OF SPEEDEDNESS IN A PLACEMENT TEST: A CALL FOR LOCAL STANDARDS

Todd M. Davis, Robert Kaiser, and Jerry Boone

Memphis State University

Jane McGuire

The State University and Community College System of Tennessee

Running head: AN INVESTIGATION OF SPEEDEDNESS

TM011761

Abstract

Higher education has struggled with the problems of extending educational opportunities to all students and assisting those with academic deficiencies by offering remedial and developmental services to underprepared students. The paradox and pitfall of these programs is that while they are designed to help students, they may contribute to resegregation of higher education when the placement procedures are not sensitive to cultural differences. The Academic Assessment and Placement Program (AAPP) is a battery of tests used for diagnostic placement of Tennessee students into remedial or developmental courses. The AAPP is a Tennessee configuration of items drawn from the College Board's multiple assessment programs and services (MAPS) item pool. This study addressed the question of the effect of time limits on the completion rate of students taking the reading comprehension section of the AAPP. The results of the analysis suggested that for the population, as a whole, and by ACT subgroups, sex, and age, no substantial effects of time limits were observed. However, important differential completion rates by race were noted at all of the 20 institutions evaluated. The trend was for whites to have higher

completion rates than blacks. At some institutions the differences were profound. The implications of these findings for the use of standardized placement instruments for remedial and developmental programs are discussed.

An Investigation of Speededness in a Placement Test: A Call for Local Standards

Since the Brown v. BOE decision, in 1954, public education has struggled with the problem of extending desegregation. Running parallel with this effort has been a concern with the academic preparation of entering college freshmen. When higher education extended remedial and developmental services to underprepared students, it also accepted the challenge to assist them in ways that alleviate the distinctions between racial groups. Otherwise, these programs might contribute to resegregation by creating separate entry tracks based on the attainment (or the lack of attainment) of basic skills. Inequitable, group-based, placement procedures might disproportionately feed a college-level track for whites and a remedial-level track for blacks or other disadvantaged groups.

The Tennessee State Board of Regents (SBR) system is committed to extending equal opportunities for higher education to all residents of Tennessee. It started a system-wide remedial/development (R/D) program in 1985. The policies that guide the program have significant impacts on the educational experiences of underprepared students. Two such consequences of R/D assessment and

placement include (1) longer time to reach the B.A. degree level and (2) greater costs in terms of tuition expenses and lost wages. These barriers, for some, could lead to dropping out. These additional time and cost considerations are not felt by well-prepared students who do not need to participate in remedial or developmental courses. With this type of differential impact built into the program, it is important that the effects of assessment and placement be felt equally among the diverse groups.

Underprepared students are applying to Tennessee's universities and colleges in large numbers. In the fall of 1986, 16,632 first-time freshmen enrolled in Tennessee State Board of Regents institutions and 8,821, or 53 percent, were placed in one or more remedial or developmental courses on the basis of their performance on required tests. Specifically related to reading, 5,838, or 66 percent of these, were placed in remedial or developmental reading courses.

Purpose of the Study

Currently, the State Board of Regents uses the Academic Assessment and Placement Program (AAPP) test to place students into remedial and developmental classes. It consists of items drawn from the Multiple Assessment Programs and Services (MAPS) item pool by the College Board. Students under 21 years of age scoring 15 or below on their American College Tests (ACT) Composite Score are required to take the AAPP tests in English, mathematics, reading comprehension and logical relationships. Students 21 years of age or older, having no ACT test on file, are also required to take the AAPP tests. If students score highly enough on the AAPP, they are permitted to enroll in regular credit-producing college courses. If not, they are placed in remedial (lower level) or developmental (higher level) courses.

The AAPP tests are designed by the College Board as timed-power tests, in which speed of work or time limits would have no significant effect. The time limits on the reading comprehension test caused concern on the part of SBR instructors and directors in R/D programs.

That initiated this study of the impact of speededness of the AAPP reading comprehension test on students at the SBR's 20 institutions.

The specific question addressed in this study was: Do the time limits imposed by the AAPP reading test foster differences in the completion rate of students with respect to their ability, age, sex, race and institution?

Effects of Speededness

There is no shortage of research on speededness; however, relatively little is known of its impact on low ability and minority students. In research involving the General Educational Development tests, Flores and Seaman (1978) found that persons who performed better under non-timed administration included persons over 39 years old and black Americans. Immerman (1980) found adult Native Americans to score higher on the Stanford Diagnostic Reading Tests under no-time-limits test conditions.

Reilly and Evans (1974), however, found that more liberal time limits did not benefit the performance of minority groups on a national aptitude test. Wild, et al (1982) found that neither verbal nor quantitative GRE scores were differentially improved by extending the

time limits for subgroups by race, sex, or years out of school. In a very relevant study, Gallagher et al. (1987) experimented with the effects of extending the time allotted for some of the subtests of Florida's College Level Academic Skills Test. They found no significant relationships among extra time, race, and test performance but commented that the sample used contained too few black students to draw reliable conclusions about race.

In a careful critique of methodological difficulties in the measurement of speededness, Rindler (1979) suggested careful attention to the possible differential group-based impacts to insure that tests are not culturally biased.

Measurement of Speededness

An annotated bibliography summarizing research on speededness has been published by Donlon (1980a), who, in a separate publication (Donlon 1980b), identifies and discusses the strengths and weaknesses of several approaches. The Educational Testing Service (ETS), which is a unit of the College Board, publishers of the MAPS and AAPP tests, has its own definition of speededness. It considers a test to be speeded if fewer than 100 percent of the candidates reach 75 percent of

the items and fewer than 80 percent of the candidates finish the test (Donlon, 1973).

Reilly and Donlon, cited in Donlon (1980b), have suggested an approach to the measurement of speededness based on the proportion of subjects who do not finish the test in the allotted time and on the number of items "Not Reached," as opposed to the number of unattempted items. The approach is derived from the ETS definition and uses a Not Reached score as the fundamental measure of speed. The number of items Not Reached is defined by a continuous string of items running from the last item completed to the end of the test. For example, in a 45-item test, such as the AAPP, if a subject fails to respond to items 40, 43, 44, and 45, the number of unattempted items would be four (4), while the number of Not Reached items would be three (3).

Method

To determine whether or not a speededness factor affected differentially the completion rate of students on the AAPP reading comprehension test, the test publisher's (ETS) definition of speededness was employed. Speededness of the test was analyzed graphically for the entire population of students who took the test during the fall of 1986 and separately by

ACT score, sex, age, race and by racial groups within the 20 institutions. Speededness was addressed through the examinations of curves obtained by plotting the percent of examinees who reached the last 30 items on the 45-item test. The ETS criteria for a timed-power test are also plotted on the graph. This graphic display technique follows that suggested by Evans and Reilly (Donlon, 1975). Specifically, the criteria against which the AAPP is compared are: (1) virtually all test takers must reach item 33 (75 percent of the items) and, (2) 80 percent or more of the test takers must reach all of the items. If both of these criteria are met, the test may be judged to be essentially unspeeded.

Results

Speededness for the SBR System

Inspection of Figure 1 suggests that for the total population of SBR students who took the AAPP reading comprehension test, no substantive indications of speededness are noted. This plot is based on the responses of 8,290 students. The mean number of items Not Reached for this total group is 1.129 with a standard deviation (SD) of 3.688.

Insert Figure 1 about here

It is the policy of some of the SBR member institutions to test all incoming students with the AAPP. This made it possible to compare the relative speededness of the AAPP for those with ACT Composite scores of 16 or greater with those who obtained ACT Composite scores of 15 and less. Figure 2 presents the comparison of these two groups with respect to the speededness of the AAPP.

Insert Figure 2 about here

Of those taking the AAPP reading comprehension test, 7,088, or 85.5 percent, earned a score of 15 or less on the ACT, while 1,202, or 14.5 percent, earned a score of 16 or greater. As Figure 2 indicates, a substantial difference between these two groups can be noted. For the higher ACT group, virtually all of the respondents completed the entire test. The mean number of Not Reached items for this group was 0.131 with an SD of 1.518 items. For the lower ACT group speed appears to be somewhat more important, but their performance does conform to the ETS guidelines with 84.5 percent

completing the test, and all but 2.7 percent completing at least 75 percent of the items. The mean number of "Not Reached" items for this lower ACT group was 1.298 with a SD of 3.914. The difference between the two ACT groups appeared to suggest that speededness was not a factor in the performance of the higher ACT group, and these 1,202 high-scoring individuals, who presumably were ineligible for either remedial or developmental placement, were excluded from subsequent analyses.

Speededness by Major Classifications

The records of 7,088 students who scored 15 or less on the ACT were analyzed separately by sex. Of this group, 4,171, or 58.8 percent, were females; 2,848, or 40.1 percent, were males; and the sex of 69 individuals was not identified. Figure 3 compares the speededness curves developed for males and females who scored less than 16.

Insert Figure 3 about here

Speed does not appear to be an important factor for either males or females and virtually no difference in the performance of the two sexes is noted. The mean number of Not Reached items by males and females was

1.296 and 1.287, respectively, with a SD of 4.080 and 3.749, respectively.

The relationship of student age and speededness of the AAPP was explored for the low ACT group by splitting the pool of test takers into two groups and plotting their completion curves separately. The older group consisted of 2,423 (34.1 percent) students who were 21 years of age and older while the larger younger group consisted of 4,665 (65.8 percent) who were under 21.

Insert Figure 4 about here

Examination of Figure 4 indicates, that for the younger group, the rates of completion appear to be satisfactory. The older students, while marginally within the ETS guidelines for a non-speeded test, appear to have a markedly lower completion rate than their younger counterparts. Among the younger group, 86.2 percent reach the last item while this is true of only 81.1 percent of older students. On the average the younger group fails to reach 1.03 items (SD=3.34) and the older group fails to reach 1.79 items (SD=4.78). As a generalization, neither group can be said to be affected by the time limits of the test. However, there

are proportionately more older students among the low ACT group who do appear to have difficulty completing the test within the time limit.

When the pool of low ACT students was split by race, 47 reported themselves as Asian, 33 as Native American and 39 as Hispanic. The 1,954 black students represented 27.5 percent of the low ACT group and the 4,946 whites accounted for 69.7 percent. As there were so few Asians, Native Americans, and Hispanics, it was felt that stable comparisons by race could be made only for blacks and whites.

Figure 5 displays a striking difference between blacks and whites with respect to their ability to complete the test.

Insert Figure 5 about here

While 88.9 percent of white students were able to complete the test and 98.5 percent, or virtually all, white students reached at least 75 percent of the AAPP items, only 73.7 percent of blacks completed the test and only 94.4 percent completed 75 percent of the items. On the average, blacks failed to reach 2.4 items (SD=5.29) while whites failed to reach .82 items

(SD=2.97). Using the ETS standard referred to earlier, speededness was an important factor for blacks who took the test, but it was not for whites.

Speededness by Institution and Race

The SBR system is composed of 20 institutions: 6 universities, 10 community colleges, and 4 technical institutes. As these three types of institutions are dissimilar in many ways, the effect of speededness was gauged comparatively within institutional type.

The records of 2,291 students enrolling in the senior institutions were examined. The test appears to function as a timed-power test for these institutions taken as a whole. But when the completion rates for these schools are examined for possible differential completion rates by race, white students have higher completion rates than black students in all of these schools.

There were differences in the completion rates between blacks and whites at all of the system's community colleges. In all cases, white students' completion rates were higher than those of black students. At seven of the nine schools examined, the differences between blacks and whites were so profound that AAPP reading comprehension performance for blacks,

when evaluated by the ETS guidelines, contained an important speed component that was not present for whites. That is, for these black students, both time limits and reading comprehension were exerting an influence on final scores on this subtest according to ETS criteria. For whites, the test was predominantly a measure of reading comprehension.

Insert Figure 6 about here

At one urban, predominantly black community college (Figure 6), 507 blacks and 142 whites were tested, with a black completion rate of 60.4 percent and a white rate of 83.1 percent.

Four technical institutes are operated by the State Board of Regents. Their results parallel that of the community colleges. At all four of the technical institutes, white students' completion rates were higher than those of blacks. At two of the schools, the differences in completion rates were profound.

Insert Figure 7 about here

At one technical institute (Figure 7), the completion rate for the 244 white students was 88.1

percent. Of the 228 black students, only 67.5 percent reached the last item. While virtually all whites at this school were able to reach at least 75 percent of the items, only 91 percent of the blacks were able to complete 75 percent of the test items.

Conclusions and Recommendations

The results suggest that of the Fall, 1986 AAPP data for the population as a whole, by ACT composite score subgroups, sex and age, no substantial indications of speededness were observed. However, the analysis demonstrated differential completion rates for blacks and whites across the SBR system, with whites having higher completion rates than blacks at all institutions. At some institutions, especially among the community colleges and technical institutes, the differences between blacks and whites were profound. When judged by the criteria developed by ETS, there is a speededness component for black students. For white students, the test functions as indicated in the test manual, namely as a timed-power test in which final scores are largely attributable to the ability to read with comprehension, not to time limits. These findings are consistent with those reported earlier by Flores and Seaman (1978) which

identified differential subgroup performances under speeded and unspeeded conditions.

It is important to note that this present study did not attempt to estimate how much impact speed requirements have on actual test scores. This would require calculation of the work rate and the efficiency with which items are answered correctly. This present study addressed only the question of whether or not the established time limit was a factor in the test results according to ETS's own standards. It was not an attempt to predict what individuals would do with additional time.

An obvious response to our finding that black students are subject to influence by a speed factor in the reading comprehension test would be to extend the time limit, so they, as well as white students, could reach ETS's standards for test completion. To do so, however, would not necessarily affect test scores and assign a different proportion of the students to the R/D program than were assigned with time limits. The way in which additional time is used in an important factor in determining final scores.

The work of Wild et al. (1982), and that by Gallagher et al. (1987), mentioned earlier, would lead to a prediction that increased time would not alter the test outcome. But the first of these studies concerned a quite different test, the GRE, and a different population; and the second was based on a sample which included very few black students.

Until this question is understood better, an R/D program which involves many black students should employ a reading test, and other tests, which conform to the publisher's own standards for power tests. This is an era when the fair use of tests for selection purposes is under close scrutiny for racial equity. Differences in how students of different races approach the testing situation should become a high priority for study. Until the time comes when we can describe fully different approaches to test taking, users of tests are urged to conduct local studies of the performance of test takers on standardized placement tests.

References

Donlon, T. P. (1975, April). Estimating rate-of-work parameters in the assessment of test speededness. Paper presented at the meeting of the National Council on Measurement in Education, Washington, D. C. (ERIC Document Reproduction Service No. ED 115 671)

Donlon, T. P. (1980a). An annotated bibliography of studies of test speededness. (Report No. GREB-76-9R). Princeton, NJ: Educational Testing Service. (ERIC Document Reproduction Service No. ED 241 565).

Donlon, T. P. (1980b). An exploratory study of the implications of test speededness. (Report No. GREB-76-9P). Princeton, NJ: Educational Testing Service. (ERIC Document Reproduction Service No. 241 560).

Donlon, T. P. (1973, November). Establishing appropriate time limits for tests. Paper presented at the meeting of the Northeast Educational Research Association, Ellenville, NY. (ERIC Document Reproduction Service No. ED 118 590).

Flores, T. B., & Seaman, D. F. (1978). A comparative study of adult student performance on timed GED tests in Texas. (ERIC Document Reproduction Service No. ED 154 138)

Gallagher, P. (1985). Time as a factor in test performance in a state mandated testing program. (ERIC Document Reproduction Service No. ED 275 729).

Immerman, M. A. (1980). The effect of eliminating time restraints on a standardized test with American Indian adults. (ERIC Document Reproduction Service Report No. ED 196 584)

Reilly, R. R., & Evans, F. R. (1974, September). The effects of time test limits on performance of culturally defined groups. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service Report No. ED 102 276).

Rindler, S. (1979). Pitfalls in assessing test speededness. Journal of Educational Measurement, 16(4), 261-269.

Wild, C., Durso, R., Rubin, D. (1982). Effect of increased test-taking on test scores by ethnic group, years out of school, and sex. Journal of Educational Measurement, 19(1), 19-28.

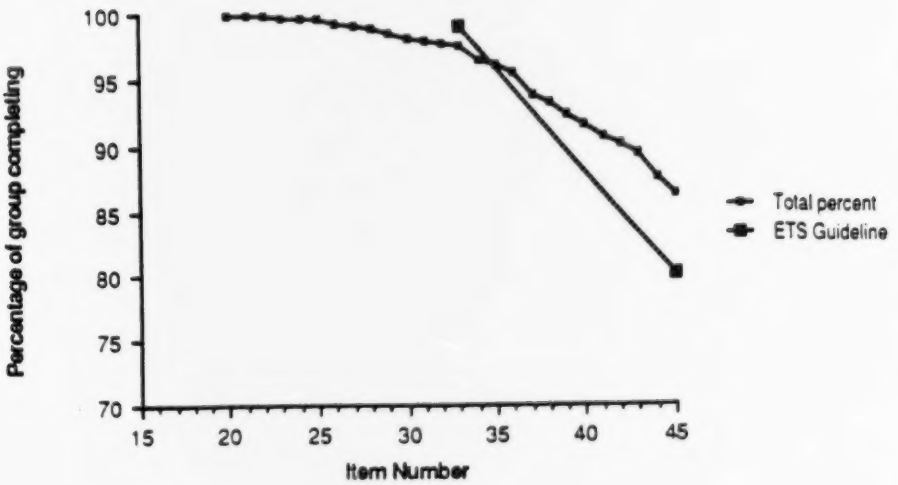

Figure 1

Items completed by total group

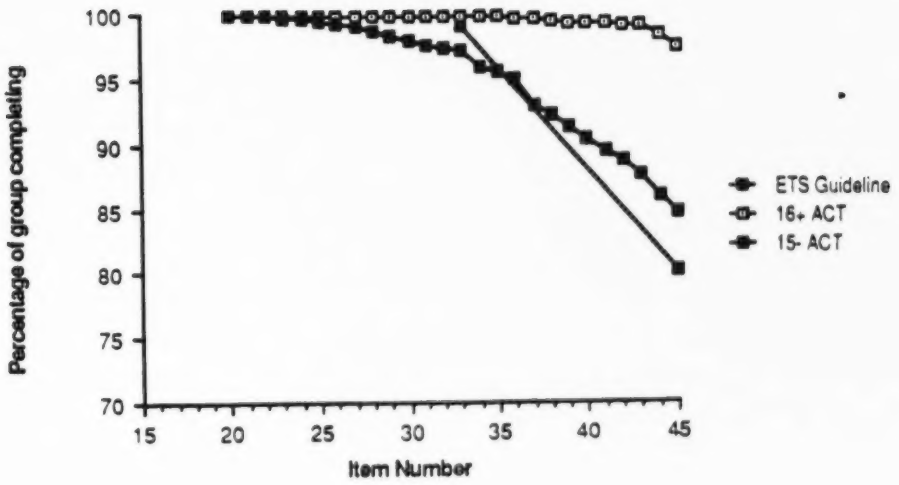

Figure 2

Items completed by ACT score

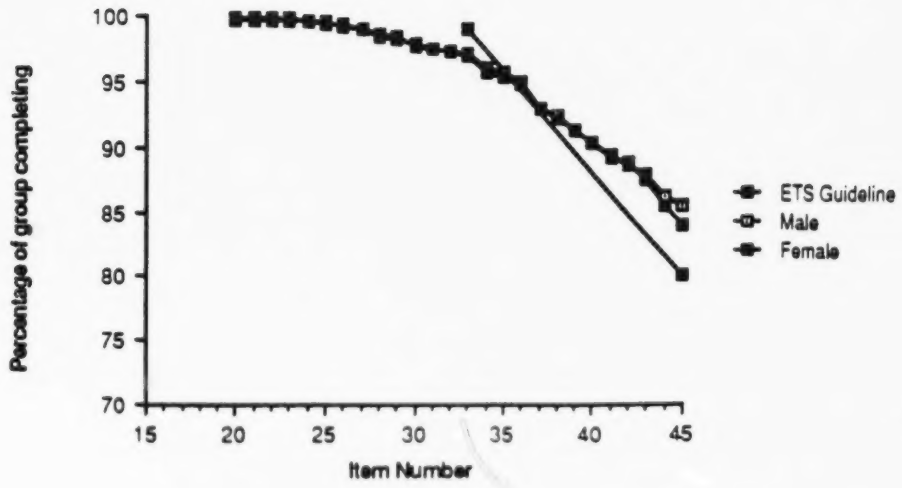

Figure 3

Items completed by sex

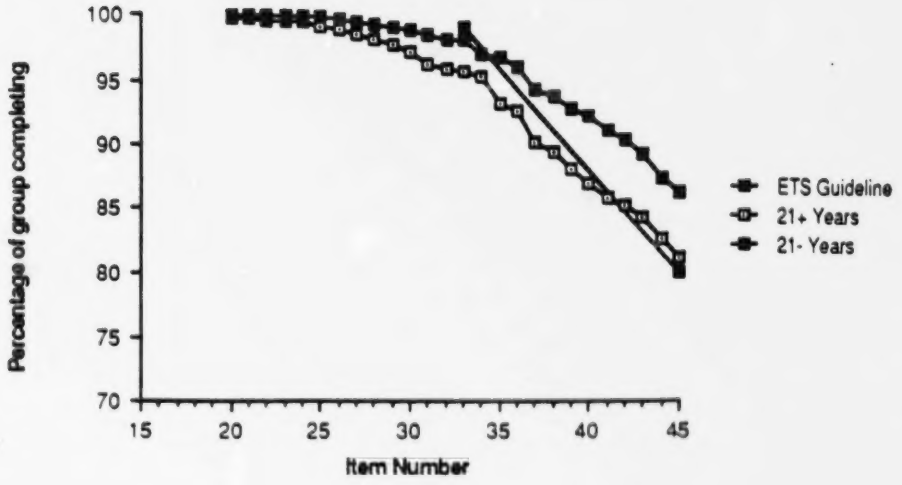

Figure 4

Items completed by age

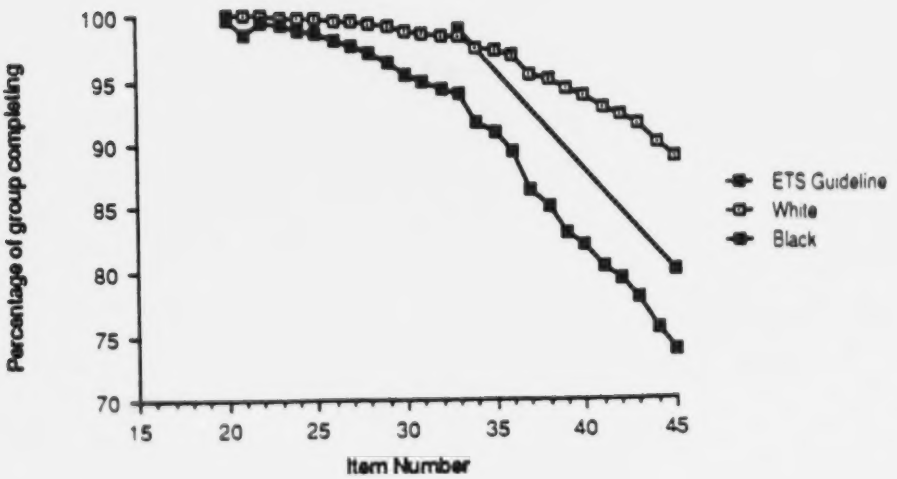

Figure 5

Items completed by race

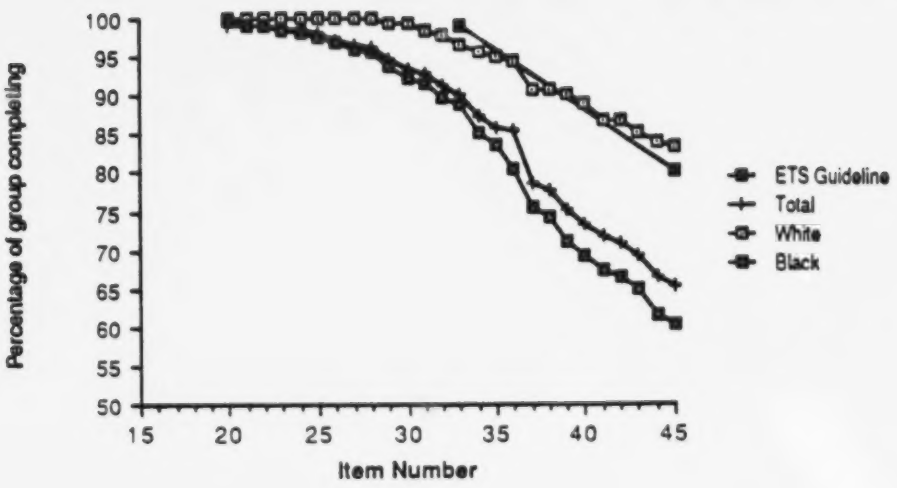

Figure 6

Items completed at an
urban community college

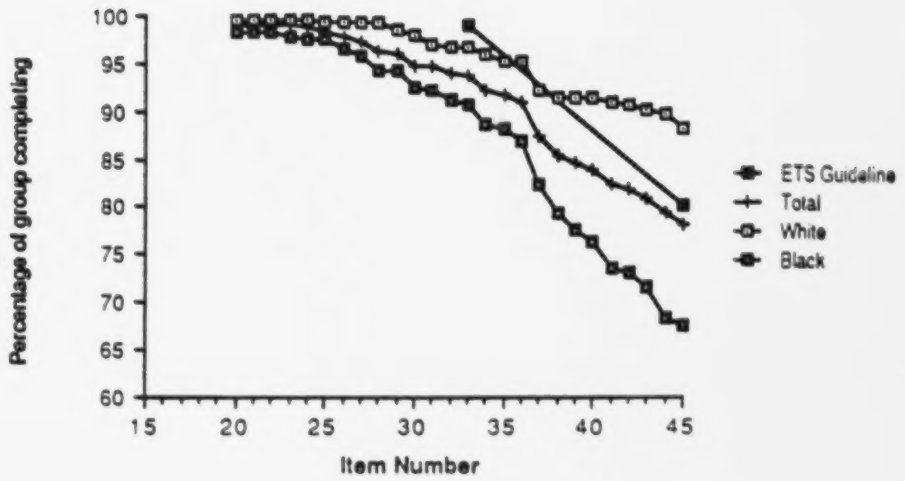

Figure 7

Items completed at a technical institute